

AN APPARATUS FOR MIGRATION AND CONVERSION OF SOFTWARE CODE FROM ANY SOURCE PLATFORM TO ANY TARGET PLATFORM

INVENTORS:

MITEN MARFATIA, SUNNYVALE CA

AJAY RAMBHIA, SUNNYVALE, CA

CROSS REFERENCE TO RELATED APPLICATIONS

[0001] This application is the National Stage of International Application No. PCT/IN2004/000385, filed December 10, 2004, published in English under PCT Article 21(2), which is incorporated by reference in its entirety.

BACKGROUND

1. FIELD

[0002] The present invention generally relates to an apparatus for migration and conversion of software code using a knowledge engine (KE) to identify the source application and to convert the same logic and database to any target application using target specific knowledgebase (KB).

2. DESCRIPTION OF THE RELATED ART

[0003] Businesses of companies, governments and agencies evolve over the years. So do government regulations. Either or both of these scenarios result in changes, deletions or addition of rules that govern the mode of doing business. Today the useful information that contains the rules of doing business (business knowledge) is embedded in the source code of

software applications and their documentation that run any organization's operations. Over the years, however, software application systems have grown beyond just the scope of maintaining system documentation and workings information. These systems have been modified, added to and deleted from as rules have changed; and the result is that they have become too rigid and hard to understand. To add to the problem, not only is the application knowledge distributed over disparate program code, but so are the data structures and data representation schemes this knowledge uses.

[0004] The problem with the various technologies available today is that they are limited in terms of their ability to discover, recognize and extract the required knowledge elements from diverse systems. The technologies are also non-flexible and produce proprietary outputs, which are not widely acceptable in the information technology (IT) world.

[0005] Another problem is adaptability. None of the available technologies are adaptable and "self learning" thereby limiting their repetitive and long-term applicability and use-ability. The prior solutions available had the following shortcomings: (1) They were unable to perform effectively transform the input data stream due to their limitations of using fixed 'static' rules; (2) They further do not ability to arrive at an abstract view for the input data stream, rather they use simple transformation and syntax matching techniques for interpreting input code on purely line-by-line basis and transforming the same; (3) The pattern recognition utilized in prior inventions or tools was not dynamic and hence not able to perform self-interpretation (i.e. interpretation on its own by dynamically looking at

parameters); (4) The previous technologies did not provide ability to dynamically (at runtime) “hatch” new patterns, which severely limited their adaptability and hence applicability.

SUMMARY

[0006] An apparatus for migration and conversion of software code from any source platform to any target platform that migrates and/or converts any source application working on any platform into a format of any target platform. The apparatus comprises an inputting means for accepting the entire source code of sample part in ASCII to analyze the business logic of the source application, obtaining UI (User Interface)/GUI (Graphical User Interface) details of the source and target application. The inputting means also receives validation schemes of source front-end interface. It also obtains the definitions of the target back-end system, the existing test scripts to facilitate the quality control phase of the generated code, the source code entry points to business processes, target environment specification or definitions which includes target platform(s), languages to be used, target database, coding standards, target architecture and framework, third party components, existing applications which have to be plugged with target application, and sample code for the application working in the target environment (if available).

[0007] The apparatus also includes an analyzing means analyzes the source schemes provided by the client to create target schemes. It also analyzes the business logic in the source application to create workflow diagrams that represent the source application

processes. In addition, it identifies the code segments in the source application and analyzes the target to generate the target architecture and the technology associated with it.

[0008] Further, the apparatus includes a setting up means for generating a custom knowledge base where the existing knowledge base is reviewed for particular migration. In cases in which no such knowledge base exists, a custom knowledge base is created.

[0009] In addition, the apparatus includes a processing means for conversion of source code in a format of a target specification, wherein the complete source code is passed through a knowledge engine (KE) on the basis of iteration. During this time the knowledge engine remains coupled to the knowledge base for conversion of source code in format of target specifications; and after each iteration the knowledge base is updated which leads to speedy and better conversion of source code as the Custom KB has now more structured information of source platform and source application with respect to target platform and target specifications.

[0010] The apparatus also includes a documenting means for generation of reports during a review of the process stage and a summary report after the end of the conversion process. The reports also include the code that is not converted automatically. This unconverted code is then converted manually at an applicant's resource center.

[0011] Also disclosed is an apparatus for migration and conversion of software code from any source platform to any target platform that migrates and/or converts any source application working on any platform into a format of any target platform comprises of: An inputting means for accepting the entire source code of sample part in ASCII to analyze the

business logic of the source application, obtaining U1 (User Interface)/GUI (Graphical User Interface) details of the source and target application, also receiving validation schemes of source front-end interface, obtaining the definitions of the target back-end system, the existing test scripts to facilitate the quality control phase of the generated code, the source code entry points to business processes, target environment specification or definitions which includes target platform(s), languages to be used, target database, coding standards, target architecture and framework, third party components, existing applications which have to be plugged with target application, and sample code for the application working in the target environment (if available); An analyzing means for analyzing the source schemes provided by the client to create target schemes, analyzing the business logic in the source application to create workflow diagrams that represent the source application processes, identifying the code segments in the source application and analyzing the target to generate the target architecture and the technology associated with it; A setting up means for generating custom knowledge base where the existing KB is reviewed for particular migration and in case of no such KB exist, a custom KB is created; A processing means for conversion of source code in format of target specification wherein the complete source code is passed through a knowledge engine on the basis of iteration and during this time the knowledge engine remains coupled to the knowledge base for conversion of source code in format of target specifications; and after each iteration the knowledge base is updated which leads to speedy and better conversion of source code as the Custom KB has now more structured information of source platform and source application with respect to target platform and target

specifications; and A documenting means for generation of reports during review of the process stage and a summary report after the end of the conversion process, which consists of the code that is not converted automatically. This unconverted code is then converted manually at applicants Resource Center.

[0012] The present invention beneficially includes an apparatus for conversion and/or migration of any source application with no limitation of platforms on which the source or target systems are working. Some embodiments advantageously allow for this by providing a unique methodology to achieve the conversion or migration activities and providing an ability to “learn” and “infer” to efficiently perform the conversion task. Some disclosed embodiments use a knowledgebase concept, such that conversion process time decreases as more source applications are processed by the applicant’s software tool. In addition, the embodiments convert and migrate using “intelligent” mechanism that utilizes neural networks and fuzzy rules. In summary it creates value by offering a unique ability to transform and convert any source to any target by using intelligent algorithms and fuzzy rules that have capability of learning to perform optimal conversion.

[0013] The features and advantages described in the specification are not all inclusive and, in particular, many additional features and advantages will be apparent to one of ordinary skill in the art in view of the drawings, specification, and claims. Moreover, it should be noted that the language used in the specification has been principally selected for readability and instructional purposes, and may not have been selected to delineate or circumscribe the inventive subject matter.

BRIEF DESCRIPTION OF DRAWINGS

[0014] The disclosed embodiments have other advantages and features which will be more readily apparent from the following detailed description and the appended claims, when taken in conjunction with the accompanying drawings, in which:

[0015] Figure 1 illustrates a block diagram of one embodiment of the system in accordance with the present invention.

[0016] Figure 2 illustrates an aided process for migration or enhancement of a software application.

DETAILED DESCRIPTION

[0017] The Figures (FIGS.) and the following description relate to preferred embodiments of the present invention by way of illustration only. It should be noted that from the following discussion, alternative embodiments of the structures and methods disclosed herein will be readily recognized as viable alternatives that may be employed without departing from the principles of the claimed invention.

[0018] Reference will now be made in detail to several embodiments, examples of which are illustrated in the accompanying figures. It is noted that wherever practicable similar or like reference numbers may be used in the figures and may indicate similar or like functionality. The figures depict embodiments of the present invention for purposes of illustration only. One skilled in the art will readily recognize from the following description

that alternative embodiments of the structures and methods illustrated herein may be employed without departing from the principles described herein.

SYSTEM CONFIGURATION

[0019] Referring first to Figure 1, it illustrates a block diagram of one embodiment of the system in accordance with the present invention. According to one embodiment of the present invention, an apparatus for migration and conversion of software code from any source platform to any target platform migrates and/or converts any source application working on any platform into a format that any target platform.

[0020] The apparatus includes an inputting means for accepting the entire source code in ASCII format to analyze the business logic of the source application. Also included is a user interface (UI)/ graphical user interface (GUI) that provides details of the source and target application. Also provided are validation schemes of a source front-end interface and definitions of the target back-end system. Existing test scripts facilitate the quality control phase of the generated code. The source code entry points to business processes and target environment specification or definitions, which include target platform(s), languages to be used, target database, coding standards, target architecture and framework, third party components, and details of existing applications that have to be plugged into the target application and sample code of the application working in target environment (if available).

[0021] The apparatus also includes an analyzing means for analyzing the source schemes provided by client to create target schemas. The means further includes analyzing

the workflow diagrams that represent the source application process, identifying the code segments in source application and analyzing the target requirements that will generate the target architecture and the technology associated with it.

[0022] In addition, the apparatus includes a setting up means for generating custom knowledge base where the existing knowledge base (KB) is reviewed for particular migration. In instances where no such knowledge base exists, a custom knowledge base is created. The knowledge base also may be referenced as a custom knowledge base.

[0023] Further, the apparatus includes a processing means for conversion of source code into format of target specification. In particular, complete source code is passed through a knowledge engine (KE) on the basis of multiple iterations. During this time the knowledge engine remains coupled to the knowledge base for conversion of source code into format of target specification. After each iteration the knowledge base is updated which results in speedy and better conversion of source code as the knowledge base gets more structured with respect to source platform, source application, target platform and target application.

[0024] In addition, the apparatus includes a documenting means for generation of reports related to process stage and end of conversion process. The generated reports also include code that is not converted automatically, the reasons for the non-conversion and suggestions on how the unconverted code can be converted manually, for example, through an applicant resource center.

[0025] Referring now to Figure 2, it illustrates an aided process for migration or enhancement of a software application. In one embodiment, the apparatus and corresponding

system is divided into two parts. The first is a knowledge base (KB) and the second is a knowledge engine (KE). One embodiment of a knowledge base is shown as block M in Figure 1. This embodiment of the knowledge base includes a relational database that is comprised of source and target code patterns and attributes. The knowledge base (element M) resembles and consolidates the source and target language syntax, control structure and programming style. The structure of the custom knowledge base (element M) is decided on the basis of format of source and target application. After each iteration the custom knowledge base (element M) is updated by the knowledge engine. This configuration helps increase the rate of automatic conversion.

[0026] The knowledge base can be accessed by standard object data base connectivity (ODBC) connection or through disk input/output (I/O) operations. The knowledge base may also involve an interface layer that will allow the knowledge engine to connect to it remotely. It is noted that element M also may be referenced as block M. However, the reference to element M helps eliminate confusion with respect to references to Blocks S, T, U, and V described in Figure 2.

[0027] One embodiment of a knowledge engine is illustrated as block N in Figure 1. The knowledge engine is a processing unit that allows the input code to be intercepted, interpreted and converted into target output code. The knowledge engine automates code migration and enhancements, allowing for many different design and development options.

[0028] To extract the business logic from the source application, the knowledge engine performs functions as described herein. In particular, the knowledge engine intercepts the

input code based on defined syntax in the custom knowledge base (element M). In addition, it interprets the source code using fuzzy logic routines. Further, the knowledge engine parses input code, segregates code blocks, and converts them into intermediate or target specific format using the knowledge base (element M). The knowledge engine also marks the unconverted or ignored blocks and sentences of the source code and stores them in the knowledge base (element M) for future use.

[0029] In case of migration of a data driven system, one embodiment of the knowledge engine operates a process as noted herein. The process interprets database schemes and data file formats and generates generate a knowledge base (block M) segment. This creates data dictionary definitions. The process interprets source code, parses it and converts it into target specific code with the help of the generated data dictionary definitions. The process marks the unconverted or ignored blocks and sentences of the source code and stores them in the knowledge base (element M) for future use. The process is repeated for arriving at the final code thereby involving intermediate processes.

[0030] Turning now to Block S in Figure 2, inputs a, b, c, d and e represent the input code components. These components are input files that are divided based on program logic code (PLC), user interface (UI), data dictionary, validation, associated program logic (APL) files or data definition files. In one embodiment, a user performs this task. Referring next to Block T, element F is a code validation stage. In this stage, the knowledge element determines if all the required input files are available for processing.

[0031] At the next stage, Block U, element N is the Knowledge Engine (KE) component of the system. This component is the primary processing module that processes the input code. Also in Block U is previously described element M, the knowledge base. As noted, the knowledge base is configured to store patterns and discovered data within the system. The symbol L in Block U represents the “iterative process” by which input code is iteratively processed by knowledge engine. The element O references a logical extension of the knowledge engine. It is configured to package the output code generated by knowledge engine.

[0032] In the next stage, Block V, element Q represents the target code generated by the tool as well as the summary report documenting the entire conversion process. Element P in Block V represents the ignored or un-converted source code. The ignored code is fed back into the knowledge base (element M) after verification to determine new patterns, which the user will insert into the knowledge base (element M) through the knowledge base configuration process.

SYSTEM OPERATION

[0033] Reference is now made to an embodiment of the workings of the system (system operation) with the help of drawings. When any source application is required to be converted and/or migrated to a target specific format then an embodiment of a process is as described herein. The first stage, referenced as an input stage or Block S in Figure 2, comprises a migration or conversion process. This stage includes receiving the source application to analyze it systematically and logically. It also includes defining the architecture

of the target application. As described herein, from this stage, either an existing library is refined or a new library knowledge base (element M) is created to fulfil the requirements of the source and target application as closely as possible.

[0034] In the input stage, a number of inputs may be received. For example, an entire source code of a sample part in ASCII (American Standard Code for Information Interchange), shown as 'a' is received so it can be analyzed to understand the business logic of the source application. From this information, the knowledge base (element M) can be defined, that is, created or updated. Reference to 'b' in this stage corresponds to receipt of user interface (UI) / graphical user interface (GUI) details of the source application. It also refers to defining the UI details of the target application. In one embodiment, for an effective conversion, validation schemes of source front-end interface are also received into this stage.

[0035] Continuing with the initial stage (Block S), the process obtains definitions of the target system database, which are referenced as 'c' in the figure. For a relational target database the scheme formats may be in structured query language (SQL) or, for mainframe computers, in fragments of the flat file database with copybook structure in ASCII format.

[0036] Next, if test scripts already exist, referenced as 'd' in the figure, these are received and used to facilitate the quality control phase of the generated target code. Reference 'e' corresponds to determining whether the source code entry points to business processes. In some embodiments, details provided by the UI may be sufficient to decide the entry points. If not, the user is asked to specify the entry points.

[0037] It is noted that the target environment specification or definitions are defined. These include target platform(s), languages to be used, target database, coding standards, target architecture and framework, third-party components, existing applications which have to be plugged with target application, and sample code for the application working in target environment (if available).

[0038] The second stage is the analysis stage, which is shown as block T in Figure 2. In the analysis stage, source code is analysed for conversion. In particular, a database analysis is undertaken in which the source schemas provided by a client are analyzed to create target schemas. A business analysis then is undertaken in which workflow diagrams that represent the source application processes are obtained. This analysis is used in generating the knowledge base (element M). The last process of this stage is a target analysis that generates the target architecture and the technology associated with it. The target architecture and the technology associated with it are added to the target application segment of the knowledge base (element M).

[0039] The third stage is a 'setup stage', which is illustrated as block U in Figure 2. In the setup stage, a generation of a custom knowledge base (element M) takes place. In particular, initially the existing knowledge base (element M) is reviewed for a particular migration. In a case of no such knowledge base (element M) exists, a custom knowledge base (element M) is created as follows herein. The selected parts of a sample code 'a', user interface 'b', and a database (back-end schemes) 'c' are introduced to the knowledge engine

(element N). The formulation done by knowledge engine (element N) on 'a', 'b' and 'c' form the base for the target requirements in the custom knowledge base (element M).

[0040] Through an iterative process, illustrated as 'L' in Figure 2, the selected source code fragment is passed multiple times through to the knowledge engine (element N). This process collects knowledge patterns. The process continues until the knowledge engine (element N) gets the signal for saturation from the knowledge base (element M). It is noted that test scripts provided by the client also are passed through the knowledge engine (element N). Alternatively, test scripts are created by the knowledge engine (element N) itself.

[0041] The next stage is the 'process stage', shown as block U in Figure 2. In this stage, the complete source code is passed through the knowledge engine (element N) on a basis of iteration. During this time the knowledge engine (element N) is coupled to the custom knowledge base (element M) for conversion of source code in the format of target specifications. After each iteration, the knowledge base (element M) is updated. This leads to a speedy and better conversion of source code as the knowledge base (element M) has now more structured information of source platform and source application with respect to target platform and target application.

[0042] The last stage is the documentation stage, illustrated as block V in Figure 2. Reports generated during the process stage are reviewed in this stage. After a conversion process, a summary report shown as P is generated. The summary report consists of code that may not have automatically converted (element O). This unconverted code is then

manually converted at an applicant's resource center. However, by this stage the applicant may have achieved 70 to 90% automatic conversion.

[0043] In this last stage two other processes are also performed. First process is a target database verification. This process includes verifying the converted database and the data dictionary with respect to the source scheme. The second process is a target application verification. This process includes verifying the converted application and its program links with respect to the source code process flow during the analysis stage (block S).

[0044] It is noted that in one or more embodiments, attributes may be formed by extracting keywords and language dependent descriptions from input supplied. The attributes may include code comments, functions, procedures, routine definition strings, parameter name strings, and main control blocks in the code for structures (e.g., if-else, do-while, while). The attributes also comprise database schemas, data file definitions and formats, and data entity definitions (including variable names, data types and size).

[0045] Also, it is noted that in one or more embodiments, business logic may be business rules for the application embedded in the input data stream. In addition, it is noted that in one or more embodiments, pattern matching may be a process by which a pattern is derived to represent a data stream to facilitate decoding information from the data stream.

[0046] In general, embodiments of the present invention described herein allow transformation of an input data stream by utilizing matching patterns and performing pattern based conversion. Further, one or more embodiments provide an ability to arrive at new patterns dynamically in case a matching pattern is not found or is not pre-defined in the

knowledge base (or knowledgebase or KB) when performing the transformation.

[0047] In one or more embodiments, the knowledge engine may use dynamic fuzzy mechanisms and rules to perform discovery, extraction, and transformation of the input data stream. In general, the knowledge engine performs the knowledge gathering process. Further, those skilled in the art will note that because the knowledge engine may use fuzzy mechanisms and rules, the knowledge engine may be “trained” to gather knowledge seamlessly across various systems and data formats.

[0048] As used herein, the terms “comprises,” “comprising,” “includes,” “including,” “has,” “having” or any other variation thereof, are intended to cover a non-exclusive inclusion. For example, a process, method, article, or apparatus that comprises a list of elements is not necessarily limited to only those elements but may include other elements not expressly listed or inherent to such process, method, article, or apparatus. Further, unless expressly stated to the contrary, “or” refers to an inclusive or and not to an exclusive or. For example, a condition A or B is satisfied by any one of the following: A is true (or present) and B is false (or not present), A is false (or not present) and B is true (or present), and both A and B are true (or present).

[0049] In addition, use of the “a” or “an” are employed to describe elements and components of the invention. This is done merely for convenience and to give a general sense of the invention. This description should be read to include one or at least one and the singular also includes the plural unless it is obvious that it is meant otherwise.

[0050] Upon reading this disclosure, those of skill in the art will appreciate still additional alternative structural and functional designs for a system and a process for an apparatus for migration and conversion of software code from any source platform to any target platform through the disclosed principles herein. Thus, while particular embodiments and applications have been illustrated and described, it is to be understood that the present invention is not limited to the precise construction and components disclosed herein and that various modifications, changes and variations which will be apparent to those skilled in the art may be made in the arrangement, operation and details of the method and apparatus of the present invention disclosed herein without departing from the spirit and scope of the invention as defined in the appended claims.